

special-topic
bulletin from

DEPARTMENT OF
EDUCATIONAL AND
PSYCHOLOGICAL RESEARCH

MALMÖ SCHOOL OF EDUCATION
LUND UNIVERSITY

didakometry

Bierschenk, B.:

AN ECOLOGICAL APPROACH TO THE PROCESSING
OF SYMBOLIC INFORMATION

No. 64

July 1981

AN ECOLOGICAL APPROACH TO THE PROCESSING OF SYMBOLIC INFORMATION

Bernhard Bierschenk

Bierschenk, B. An ecological approach to the processing of symbolic information. Didakometry (Malmö, Sweden: Department of Educational and Psychological Research), No. 64, 1981.

This article attempts to prove that an adequate cognition oriented approach to the processing of symbolic information abstracted from verbal expressions must consider a running text instead of scattered sentences. It is argued that the dominant cognition models do not give due consideration to the fact that a valid abstraction of information structures has to be based on an explicit encoding of intentionality and valuation. Following a critical discussion of the foundations of the existing models, it is stated that a model is required that can cope with empirical context and novelty instead of truth-values in semantic-logical contexts. The proposed theory has been validated on empirical text material considered to be anamorphic. It is shown that its underlying key relations can be extracted by means of discriminant functions and given a meaningful interpretation.

Keywords: Abstraction, anamorphic text, conceptual functions, contextual information, perspective, symbolic information processing, specification, structural analysis.

CONTENTS

Acknowledgement	3
1. INTRODUCTION	4
2. FOUNDATIONS OF COGNITIVE MODELS	7
2.1 The Logic of Organ	8
2.2 The Logic of Action	10
2.3 The Logic of Conception	12
3. PROCESSING AND REPRESENTATION OF INFORMATION	
3.1 Assessment of Symbolic Information Structure	19
3.2 Empirical Valuation of Conceptualizations	22
3.3 Representation of Discriminant Functions	26
4. CONCLUSIONS AND SOME VALIDATING REMARKS	30
5. REFERENCES	33

Acknowledgement

This article is based on many key ideas and studies which have been developed within several research programmes financed by the National Board of Education. These programmes would not have come into existence without its support. Within these programmes many people have worked together in order to make a systematic study of symbolic information in anamorphic text approachable.

First of all I would like to express my gratitude to my fellow researchers who without reservation have served as experimental subjects. It is a pleasure for me to convey my appreciation to Dr. Inger Bierschenk, for her discerning contributions in making an extensive text accessible to computer-based analyses. My gratitude also includes the research students and research assistants who over the years participated in certain phases of the programmes. In this connection, I would like to thank especially B. A.s M.-L. Annerblom and M. Berg. I would like to take this opportunity of thanking particularly B. A.s L. Robertsson and A. Sternerup-Hansson for their contributions to the development of several computer programmes and for carrying out the individual analysis programmes.

1. INTRODUCTION

The study of cognition requires methods and techniques for an abstraction of information from verbally expressed experiences. Only when people can express themselves verbally, i.e. freely and unrestrictedly, does it become possible to get access to information of high validity. The analytical problem in making inferences from verbal expressions is much the same as with other types of data. Specific events, behaviors or qualities characterizing certain objects and events have to be inferred. Thus, the process of inference making requires that relationships between events and verbally reported experiences can be established.

The immaterial character of the concept of information is further emphasized through the integration of computers in the technically advanced societies. "Data flow" and "transformation of information" are descriptive, although abstract, concepts, labelling processes whose operation is difficult for people outside the programmer circle to conceive. The development of electronic "information processing machines" has started a new phase by the abstraction of information from symbols. At the same time one may ask whether, and the extent to which, information from psychological phenomena can be abstracted and represented symbolically through the logical formalism that forms the basis of computer technology.

Symbols may be considered as holding abstracted information. But the ability to abstract information from strings of symbols presupposes not only a device that has been provided with the capacity to "retrieve" a string of symbols, e.g. a verbal expression, but also that there has to be a means of "storing" of context. Without access to context, symbols cannot be recognized and used in the processing of symbolic information, since the risk of misuse is too obvious.

The emphasis of this article is placed on a theoretical discussion as well as on the empirical study of information structures contained in anamorphic text as opposed to well-formed but scattered sentence examples cited by generations of philosophers, linguists and psychologists. The goal is to present an ecological model for the processing of information abstracted from symbols. This requires that the proposed model can cope with empirical context, and novelty.

From a cognitive point of view information derived from symbols does not seem to be easily separated from its social context. Therefore, information may be examined with respect to its form (form of symbolism) on the one hand and its variability (degree of abstraction) on the other.

The model to be presented here is based on the assumption that human cognition does not follow the same strict formalism as that of electronic information processing machines. The psychological approach taken stresses the fact that humans ordinarily can recognize objects and events with absolute certainty. Despite this fact cognition oriented scientists and some quarters in cognitive psychology stress the formal-logical approach, and describe the phenomenon as "expectation" or "hypothesis", because computers function according to principles of formal logic and therefore show a logical behavior.

A context dependent analysis of information requires a decision regarding who or what should represent the field or context. The model to be presented was developed in connection with the analysis of an extensive interview material concerning social science researchers' grant-supported activities. Based on the results presented in B. Bierschenk (1974) the definition of a population of researchers employed here includes psychologists, educators and sociologists. After "researcher" had been defined a strictly random sample was drawn (32 %) from the resulting

population (N = 126). They served without reservation as experimental subjects in the investigations that generated the text material.

The information contained in an interview text and the way its organizational set-up is laid out decides what is recognizable and what is not. The interview text can be considered to be characterized by cues that are ambiguously organized to a certain degree. An analysis of the interview text requires that information about the viewpoints that are at different locations has to be combined into a single conceptualization. In the case where several viewpoints of central importance for the cognition of a certain event fall outside a particular perspective, the attempted conceptualization remains fragmentary.

Thus, when a certain passage of text has to be processed, all the cues to the information necessary for the conceptualization to be formed may not be present in that particular part of the text. Within the study of perception, anamorphically displayed figures have been produced with different optic perspectives (e.g. magnification along mutually perpendicular radii). The analysis of perception of anamorphic figures helps to find out how the figure is normally perceived.

Empirical evidence (Kennedy, 1980) points to the fact that a viewpoint can be chosen, and the distribution of informative cues be arranged to fit with the chosen perspective. Some empirical results of the study of visually disabled children show that perspective not only pertains to perception (vision) but also has cognitive properties.

With respect to the study of information in verbal expressions, the same principle seems to have been applied by Jenkins (1973) who could show that his subjects made use of separately presented verbal cues to reconstruct the unity of an event. The cues may appear in a text that is composed by various linguistic rules.

Thus, the point of view or the perspective chosen may not be the most fundamental property of a model for the analysis of symbolic information processing, not even for the analysis of perceptual information processing (cf. Kennedy, 1980). Anamorphically organized pictures and verbal expressions point to the circumstance that information can be preserved while the perspectives change.

2. FOUNDATIONS OF COGNITION MODELS

Within the framework of human cognition two general hypotheses can be discerned, namely (1) the frame hypothesis and (2) the schema hypothesis (see B. Bierschenk, 1980). The first assumes that perceptual constancy has to be established before any account of cognition (symbolic information processing) can be given. The second hypothesis has advantages with respect to the completeness of the information to be processed, i.e. the preciseness of the cognition of objects and events can vary widely. What type of assumptions are made with respect to the logic of the (1) organ, (2) action, and (3) conception (of Piaget, 1978) will be discussed below.

The notion "information" is central to the models and theories based on the assumptions connected with both hypotheses. Because the study of cognition has been influenced by many disciplines, there have been major shifts of emphasis. Cognition is, as a rule, considered to develop in a linear fashion out of the "wetware" of the organ or the "hardware" of the machine. Therefore it is believed that a theory of cognition has to incorporate the theories and experimental results concerning information processing within the lowest possible level of the organism or the machine.

Starting at the physiological bottom, it is believed that the infor-

mation processing at this level somehow converges at the highest level into symbols ("human knowledge"). Wolsky & Wolsky (1976, p. 106), for example, formulate their position as follows:

"A bee, being aware of its own position in the environment, experiencing distance and quantity (of food) and translating these experiences into abstract communication (body, language dance), must be to a certain degree self-conscious. If one does not assume that consciousness appears suddenly and in a very unique fashion only with man, one has to admit that human consciousness was in the coming."

It is obvious from this quotation that the authors locate "cognition" in the structure of behavior. They seem to suggest that the animal has some ability to develop symbols, i.e. an understanding of the meaning of "self", because its lowest common denominator is suggested to be the relation between the organism or "agent" and the object of his action, in which agent and object reside in the same organism.

2.1 The Logic of Organ

Under the frame hypothesis it is expected that a visual pathway from the eye to the cortex breaks a retinal image down into a number of abstract features (cf. Regan & Beverley, 1977). It is assumed that the passage of time is homogeneous and linear and can therefore be used to investigate information processing by means of reaction time measures. Further, it is also assumed that distance in the wetware can be conceived as existing invisibly and therefore can essentially be modelled as an abstract space. Founded on these assumptions, stages of memory (see Sperling, 1963) are postulated as the adequate theoretical conception for the study of information processing. But Di Lollo (1980) has attempted to question the most widely accepted hypothesis of the existence of an Iconic Memory. His results do not support this hypothesis. A theory of

parallel functioning information processing channels is clearly more favored by his results.

Sperling had already (1963) found some empirical evidence against the processing of information within a system of memories. He concluded that information does not exist in any form that could be made objective (conscious). Furthermore, Wickelgren (1975) in a study on the rate of forgetting as a function of age, tested cognition instead of the usual recall of information. The results of the study point to the fact that the "retention functions" for different age groups essentially are the same. This places some confidence on the hypothesis of "formless invariants over time" as is suggested by Gibson (1966, 1979).

According to Gibson's view information is picked up by the organism in order to transform it. As Gibson (1979, p. 258) notes the individual extracts and abstracts invariants in both the processes perceiving and knowing. Regan, Beverley and Cynader (1979) have developed a psychophysical model of stereoscopic motion. Results founded on this model seem to support Gibson's assumption of "super-ordinate components" and "invariants" existing in the visual flow pattern.

Thus, these results imply that cognition cannot easily be accounted for by some "pre-programmed relationship", even though Hubel and Wiesel (1963) in their contour coding experiment could show that young kittens without visual exposure have essentially the same elementary physiological contour coding system as adult cats.

Whatever caused the experimental subjects to take lines equivalent to "ordered gestalts" (edges of a surface), this is not to be considered as a "structural transformation" in order to use information intentionally. What the experiment shows is that elementary processes such as "condi-

tional judgement" (if ... then anticipation) "classification" (generalization) or "comparison" (combination) are essentially unlearned activities. Hubel and Wiesel's experimental results can be accounted for by the frame hypothesis, because this is not cognition (or learning) with an orientation toward novel solutions.

2.2 The Logic of Action

The method of association of "primitives" or privilege of co-occurrence is perhaps the most widely known approach in the analysis of behavior. Primitives are defined as analyzed entities that need no further explanation. They can be extracted from objects or events and processed directly.

Working from the frame hypothesis, gestalt psychologists have taken the outline depicted in a visual display as being a surrogate of reality. The classical bases for these studies are "size", "distance" and "form" (cf. Uttal, 1973; Winston, 1975). Furthermore, it is assumed that a "standard context" would provide for a complete knowledge of the adequate action.

The function of line drawings in caves seem to have much in common with the "figure-ground" assumption of the gestalt psychologists. Kennedy (1980) assumes that line depiction of depth change and slant change are basic perceptual phenomena arising without prior exercises. He shows (p. 16) that perspective not only pertains to vision but has cognitive properties and notes that "... perspective is in the perceptual system of the blind". Thus, perceptual experience is assumed to be direct. Moreover, a figure outlined in a visual display is considered to contain a perspective. Therefore, a description of objects and events puts both optical as well as perspective invariants into verbal expressions.

To get access to both the facts encoded symbolically and the viewpoint from which these are conceived, requires a method that tears apart the invariants of the environment from the invariants of the perspective.

Jenkins (1973) presents some experiments which were designed to investigate the importance of the perception of events in the processing of verbal expressions. The experimental conditions were set up to prove the hypothesis that the cognition of verbal expressions depends on the information that can be abstracted from an event. The tasks were constructed so that each group of expressions contained "verbal cues" belonging to an overall event.

The result of the experiment seems to indicate that the quality of each of the individual events constitutes the total information of the experimental material (a complex sentence). Changing the tasks means, according to Jenkins, a change in the kind of event that subjects in the experiments reported. Furthermore, if the overall event is available, it is possible for the individual to reject or accept verbal expressions as likely to have occurred, even though the exact wording of the expression that actually occurred cannot be reported.

Thus, concepts seem to be recognized on the basis of described events. What matters are the invariants of objects and events that are "constitutive" when the individual comes into contact with their language encoding. Thus, the individual seems also to reconstruct the structure of an event mapped onto a certain verbal expression.

The assumptions underlying Jenkins' experiments presuppose that cognitive structures develop as a result of "progressive actions" within a structurally defined context. His results imply that cognition changes continuously, partly dependent on the information processing strategies

employed, partly dependent on the kind of information available. "Novel recombinations" make possible progressive changes in the information processing strategies, which in turn cause experiences different not only in quantity but also in quality.

However, the effects of experience are not looked upon as learning in the classical sense. It is the effects of exercises that determine the individual's cognitive structure, not the contrary. The identity or unity of a cognitive structure undergoing change relies on structural invariants, which can be detected.

2.3 The Logic of Conception

The difficulty in keeping apart intention from value seems to be tied to the fact that every meaningful behavior combined them, i.e. interaction with the environment seems to require the ability to express values behaviorally.

One way of conceptualizing the structure of cognition is by means of feature models. In accordance with these models a behavioral expression is analyzed into its various primitives which comprise the lowest possible level of analysis. The primitives are then "compared" with respect to so-called features. A conceptual approach to the operational definition of concepts based on the detection of characteristic features is Quillian's (1968) "semantic network". Features of so-called semantic primitives are used as the basis for the identification of concepts. Word lists are used as formalized arrangements and are given values according to various ways of assessing their meaning. The values are then used as the basis for identifying conceptual patterns.

Recent trends have been characterized by grouping "related" primitives into various types of topical hierarchies, chunks or slots (see e.g.

Broadbent & Broadbent, 1978). The approach relies on "conditional judgement" and is intended to establish an average member belonging to a semantic-logical class or topic.

If cluster analysis techniques are used, it becomes possible to determine clusters on an intersubjective basis. In addition, within a certain analysis definable and exactly comparable boundaries can be drawn between all clusters (topics). If a criterion has been established, the manifest structure within a cluster arrangement becomes completely dependent on the value of the coefficients and thus is no longer exposed to subjective manipulation.

In a cluster analysis, each fusion criterion presupposes certain given mathematical assumptions and these vary in different cluster analysis models. By applying several different models and criteria it becomes possible to decide to what extent a certain topical pattern reflects an artificial structure or a valid pattern of topics.

An alternative approach to the analysis of verbal expressions is represented by the so-called template models. These assume that cues can be identified in lower domains and used to signal parts of a verbal expression to be transformed into a canonical representation to some higher level of information processing. Probably the most widely known model of this kind is Schank's (1972, 1973, 1975) "Conceptual Dependency Relations" approach.

Schank has developed a so-called C-diagram ("Conceptual dependency network") to express formally defined dependency relations that exist between two categories of concepts - independent and dependent. This terminology is clearly anchored in Quillian's (1968, p. 232) memory model, where the parameters S and D represent the independent concepts (subject, direct object) while the parameter M refers to dependent con-

cepts, i.e. modifiers. Both Quillian and Schank remove all inflections during the encoding process and thus transform all words into a more or less abstract canonical representation.

Schank's terminology seems to be influenced by image theories (e.g. Paivio, 1971). From a linguistic point of view he seems to believe that syntactic categories such as nouns, adjectives, verbs and adverbs reflect images. From a psychological viewpoint, Schank's encoding system implies that cognition is based on images. These are considered to be representations that correspond directly with concrete things.

When an image is evoked by a word or clause, information processing occurs, where the processing refers to the activation of structural links describing the image. According to the template principle, the most homogeneous fragment of a verbal expression is accepted.

Schank's model implies that information represented in a sentence may be retrieved by the arousal of the image, which may be the case for concrete words but applies in far lesser degree to abstract words (cf. Jenkins, 1973; Paivio et al., 1975; Tulving, 1972).

What Schank (1972) calls "Conceptual Dependency" would better have been labelled "Image Dependency". A further argument that underlines the inappropriate reference to "symbolic" structure is that Schank in fact builds on Fillmore's (1968) case structure model which is based on philosophical and not on psychological assumptions. In addition, Schank's labelling of "primitive acts" leads one to believe that he builds on "semantic primitives" suitable for manipulation (e.g. by the computer) without further explanations. On the contrary, he performs "classifications" of actions. Therefore, this seems an unfortunate misnomer. It would have been far better to label them "general acts".

Finally, Schank proposes that events comprised in episodes are the

basis for an organization of symbolic structure. Consequently he conceives his model as being able to account for an integration of events, i.e. events in which a change of state is ascribed to an object (cf. Schank & Abelson, 1977). Since his model is based on the frame hypothesis it works on a holistic basis with "ordered gestalts" and "conditional judgements" but no provision is made for the consideration of a sequential order of events and integrative functions. All that can be expressed are, at most, states, and state changes.

Thus, what can be expressed by Schank's model is something quite different from progressive action and change. The change in the representational system remains to be inferred (see Bierschenk & Bierschenk, 1976). It is important to recognize that his model does not identify the directiveness of the relation between agent and objective. Therefore, all that can be accomplished is a mediation of changes in certain attributes, e.g. a change in location. Although relying on different arguments, Cercone (1978, p. 19) comes to the same conclusion when he states:

"... Schank and Wilks deny us the option of regarding certain actions as having an agent."

The notions concept and conceptualization label the latent aspects of a verbal expression. In addition, abstracting transformational invariants from a text and developing knowledge will be considered processes that does not possess the manifest and mechanical properties assumed by the frame hypothesis. Psychological processes that create symbols are considered to be quite different from processes that imply a copy that can be stored, observed or retrieved from the brain (cf. Pylyshyn, 1973).

A way of approaching verbal expression with respect to its symbolic

nature is represented by the Agent-action-Object(ive) (AaO) model. Werner and Kaplan (1963) noted that the AaO model is basic for an analysis of symbols. A study of symbols requires that information to be abstracted is characterized as context dependent. This is no disadvantage with respect to its degree of precision and completeness, because the schema hypothesis provides the basis for the definition of context and the agent within that context.

The model is fundamental to Indo-European languages and therefore should easily be recognized as the adequate level of generality for the definition of cognitive functions. According to Werner and Kaplan (1963, p. 57) the AaO model is used for

"... connoting states of affairs and articulating them linguistically to formulate attributive propositions and propositions about conceptual relations in terms of that action model."

Thus "integration" and "unconditional judgement" can be formalized through a utilization of the AaO model. This means that symbolic information contained in verbal expressions cannot be reconstructed from individual words only, even though people often communicate knowledge or thought as economically as possible, i.e. in a fragmentary way. Therefore a first measure to be taken in an analysis of concepts and conceptual relations is to establish a context through the AaO model. It is quite obvious that the logic of conception based on the schema hypothesis contrasts favorably to considerations based on formal logical operations such as class inclusion, denotation and context independence.

3. PROCESSING AND REPRESENTATION OF INFORMATION

Organizational variables account for the richness of variation in the manifestation of one and the same conceptualization in various verbal

expressions. Moreover, these are characterized by certain parts being mobile, while others have fixed positions. Conceptualizations may manifest themselves in simple or more complex patterns of expression. Thus, an expression consists of syntactic units, which means that the syntactic position of a word can alter its informative value. Syntax implies sequence or a manifest relation between the various parts of an expression. Therefore, different elements of language can be used to reconstruct a concept such as "agent" or "object(ive)" and the relations obtaining between concepts.

The component Agent in the AaO model is used in the sense of "action centre" no matter how easily it is identified. Consequently certain events just happening are regarded as instigated by something including natural forces, i.e. the laws of nature. As already indicated, the great majority of analyses of verbal expressions are all based on the study of isolated words or isolated concept complexes. However, any sensible analysis of verbal expressions has to take into consideration the whole AaO model, i.e., the agent's coincidences with the object(ive) and the reverse. In such an analysis coincidences are to be determined by the verb linkage.

In order to encode a conceptualization, a system was developed that is presented in Bierschenk & Bierschenk (1976). The system is built up in such a way that each concept is specified by means of a two-figure code. AaO is expressed as 30+40+50. Code 30 denotes the agent, 40 the action and 50 the object(ive) function. It is not stated whether the agent is a person or an abstract concept.

An example may illustrate how an encoding according to the AaO model relates and extends over and above Schank's (1973) model. The expression which is used by Schank in order to show how his theory works is

John hits his (John's) little dog (1)

30 40 51 52 50

The conceptualization (1) consists of an intentionality dimension which has a code number ending with zero and an extensional dimension which is encoded by a final number other than zero.

One begins with the hypothesis that nouns represent labels which are used to denote abstractions. These are called "extensionals". When all the properties presented by a noun are connoted, no further analysis is possible. On the other hand explicitly expressed common-noun intensions can be analyzed further. This way of reasoning led to the construction of a coding system (see I. Bierschenk, 1977), where extensionality is differentiated from intentionality.

As shown by the coding example (1) a major difference is that the object is not coded as belonging to the verb. The coding gives expression to a dynamic model which not only gives due consideration to facts but also reflects intentionality. This makes it possible to establish a symmetrical relation between the agent and object complex. Thus, the object has the same basic structure as the agent from the point of view of their composition of main words and various attributes attachable.

Schank (1973) on the other hand codes dependency by a two-directional arrow between the "independent concept" "John" and "hit". The dependency between the object "dog" and the action "hit" is represented by a one-directional arrow from "dog" to "hit". The "dependent" concept, i.e. the conceptual attribute "little" which is predicated with respect to the "independent" concept dog is marked by a one-directional vertical arrow from "little" to "dog". Finally, the expression (1) also has an element

for possession (his). The concept dog that is attributively differentiated with respect to "John" is encoded by a double lined and one-directional arrow from "dog" to "John" which is labeled "Poss-By". The example (1) shows that the links between elements of language have been encoded multidirectionally, since several dependencies are considered. Thus Schank (1973) does not encode intentionality but does give expression to formally defined dependency relations of a semantic-logical type and refers to "standard context".

3.1 Assessment of Symbolic Information Structure

An investigatory strategy suitable for an analysis of ill-defined problems is represented by the interview method. Under the assumption that information search problems emerging during the development of a research process constitute especially ill-defined problems, the following hypothesis can be formulated:

Hypothesis 1

Researchers' free and uncommitted comments on information search problems provide symbolic information of high validity, though organized within an extremely anamorphic textual composition.

What will be abstracted from such a text depends on what is being recognized as its informative parts. Studies of the appearance of a text which results from interviews show that the recorded and typewritten text is characterized by its informative parts being distributed throughout the whole text material.

Based on the coding according to ANACONDA (Analysis of Concepts by Data Processing) the relation patterns between the agent and object(ive)

components generate very hollow data matrices (see B. Bierschenk, 1977). By using cluster analysis models it is possible to condense both sides, so that the matrices become more complete, i.e. a marking for occurrence in each individual cell.

The analysis strategy has been carried out in order to show that symbolic information structures in anamorphic compositions of text can be transformed into a meaningful and consistent information structure, i.e. a metamorphic structure. The text material consists of approximately 4 000 pages. But the answers to be analyzed make up 10 % of the material. These 400 pages of typewritten text refer to answers to the following four questions:

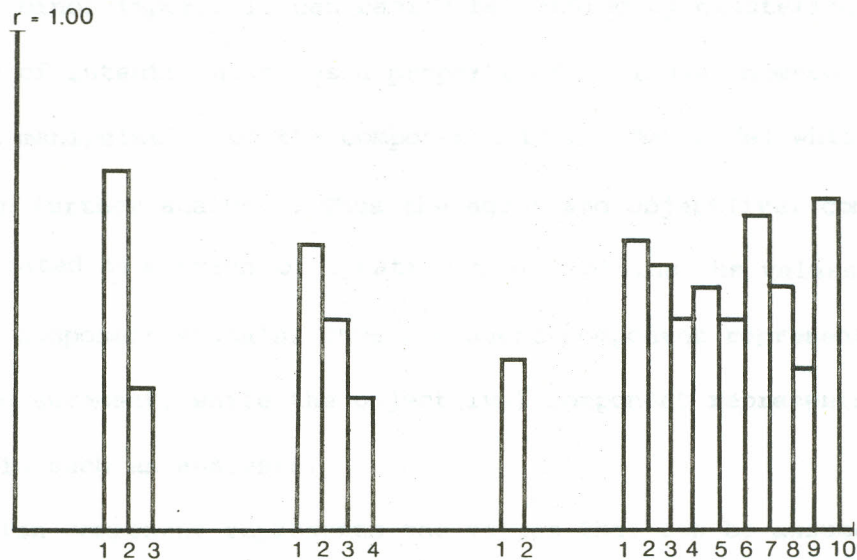
1. In which way have you tried to gain more detailed complementary knowledge?
2. How consistently during the formulation process have you made use of channels of information such as libraries etc.?
3. What type of information have you searched for and which search strategy have you used?
4. Could you say anything about how one should design information searching in order to create ideal conditions for the research process? Have you any suggestion for improvement?

All four questions concern information problems. This topic of the interview was deliberately chosen to provide for a later interpretation free of contradictions.

The coding procedure produced ca. 35 000 units that were transferred to IBM cards. An iterative application of the clustering technique chosen condensed this heterogeneous material considerably. The measurement used for the association between pairs of agents and pairs of object(ives)

respectively is "Euclidean distance". This distance is defined as "the square root of the sum of the squared differences between the values for pairs of variables". The amalgamation algorithm used in the analysis is "the average linkage algorithm". The rule states that the mean distance is to be calculated between a variable in the first group formed and a variable in the second. The similarity coefficients have then been transformed into product-moment correlations.

The summary of the agent structure appears to express agents taking contact or allowing for contact while the object(ive) structure expresses more the actual research work and means of approach. Two examples of the agent and object(ive) clusters from the final cluster analysis are given in Figure 1.



Agent Information Dissemination

1. Literature
2. Symposium
3. Documentation

Agent Methodology Oriented Researchers

1. I (a) together with area of interests
2. I (b) together with area of interests
3. I (c) together with area of interests
4. I (d) together with area of interests

Object(ive) Problem Discussion

1. Problem
2. Discussion

Object(ive) Information Channels

- | | |
|----------------------------|-----------------------|
| 1. Symposium | 6. Department Library |
| 2. Person | 7. University Library |
| 3. Library | 8. ERIC |
| 4. Psychological Abstracts | 9. Handbooks |
| 5. Reference Group | 10. Reviewing Organs |

Figure 1. Fragments of the Dendrograms reported in B. Bierschenk (1976).

The values of similarity coefficients can vary between 1.00 (perfect agreement) and .00 (no agreement at all). The lowest limit used to merge variables into a cluster formation as exemplified is a similarity value of .30 expressed as a product-moment correlation. From the point of view of interpretation, the clusters appear meaningful, but it is still premature to try to make a final judgment on the information structure underlying the interview text. On the other hand the clusters provide the necessary conditions for exploring the dimensionality of its information structure.

To separate object(ives) into different categories seems to be a problem of minor import. It can easily be handled by clustering techniques. The coding of intentionality as a property of a clause, however, is central for matrix manipulation of the components in the AaO model which will be required in further analyses. Thus the agent and object(ive) components will be treated as entries of a matrix that contains the values that the action component signals. Thus the agent component represents the units of measurement, while the object(ive) component represents the variables in such an analysis.

The action component represents the values that can be assumed by the variables attachable to that component. By linking the values of the action components with the object(ive) cluster, each cluster is provided with information values and thus a matrix can be formed which makes possible the estimation of the dimensionality in symbolic information structures.

3.2 Empirical Valuation of Conceptualizations

The discussion so far points to the central role of the verb in the

conceptualization process. Based on the assumption that it is the verb that gives informative value to the AaO model, all verbs have been extracted from the text in order to determine the degree of activity, intensity and evaluation expressed by the text. A distinction is thereby made between what constitutes a concept and how that concept is valued.

Nouns combined with their respective verb within a certain expression constitute a concept. The following example may illustrate this procedure:

I have contact with the editor (2)

30 40 50

The combination of "contact with the editor" would be considered as expression of the "fact" or concept "editor contact". The building up of concepts requires an empirical context, and presupposes a system of rules that states how and in what order different parts are to be linked together, which agents coincide with established concepts and how often they are to be determined by verb linkages. Thus a certain verb is tested with respect to all existing agent-object(ive) combinations involving that verb.

If the concepts derived from a certain text can be quantified by means of the values representing the covariation of conceptual properties, it is possible to determine functions and the weights that each property should be given in the prediction of concepts and conceptualizations.

From the interview text 882 verbs have been extracted and were changed so that the infinitive form represented different variants. The verbs were treated in this way in order to create files for deriving scaled dimensions.

A method often used for studying the dimensionality in meaning is Osgood's Semantic Differentials. Osgood (1969) claims that the established

three scale dimensions "Evaluation" (E), "Activity" (A) and "Potency" (P), or dominance, are in fundamental agreement with Wundt's (1918) three ways of classifying value. The highest loading pairs of adjectives for the scaled dimensions are respectively:

Positive-Negative	equals	Pleasure-Listlessness
Active-Passive	equals	Stress-Detachment
Strong-Weak	equals	Excitement-Calming

Semantic differentials have usually been used to describe the dimensionality of a certain noun by a set of adjectives.

The approach taken for the scaling of the verbs differs markedly from the classical scaling procedure. The scaling strategy chosen circumvents the problem that the semantic structure in the selected adjective scales changes as a function of the noun to be assessed, instead this fact is treated as an asset.

Under the assumption that a concept can be valued only to the extent that its properties have been stated explicitly by the individual researcher himself, the verbs have been scaled out of context by means of the three pairs of adjectives given. Thus, different verbs with different values are involved in the valuation of concept clusters and the analysis of symbolic information structures. A detailed account of the experimental procedures employed is given in Bierschenk & Bierschenk (1976).

It is assumed that the researcher's observations concerning himself and his environment are validly reflected by his verbal expressions. Therefore, it should be possible to distinguish informative from non-informative parts in anamorphic text. On the other hand, language is, unlike "line drawings in cave art" or abstract figures such as like triangles, considered to be a product of conventional agreement. This implies that each of a multitude of aspects may exert a significant

influence on the informative value of certain parts of a verbal expression. If an expression is informative, this should manifest itself in a value structure that is invariant while perspectives change.

Then, by definition, a certain individual possesses certain perspectives that may differ from any other individual. In order to represent different perspectives a panel of fifteen randomly selected researchers from the population defined (cf. section 1) was set up to assess the values of the verbs.

With the purpose of studying the dimensionality in the value structure represented by the verbs, a reduced component analysis was carried out. The method assumes that the values a judge assigns any one of the verbs are not determined by only one source of variation (i.e. one perspective) but by several.

Thus, if the fifteen judges of the panel basically value the verbs in the same way this should give rise to one common component. However, if two or more components emerge, this would imply that different perspectives exist, depending on how the judges have weighted on each dimension. Here it will suffice to state that the valuation has given rise to one common component in each set of scores. The maximal reliability of the values has been calculated from the largest characteristic root of the correlation matrix. The estimations for the E-, A-, and P-scaled dimension are $\alpha_{\max} = .95, .93, .86$, respectively.

Having established the information value carried by the verbs, a weighted summation variable has been created from the first unrotated component. An example of the value assigned to a verbal expression within the AaO model may show how the values of a verb are linked to its objective:

Agent	action	Object(ive)	(3)
		E 5.07	
ARBASS	is responsible for	A 4.13	literature
(Research project)		P 4.67	
		E 1.87	
Booksellers	fob off	A 6.40	books
		P 5.13	

The example (3) gives the three different values that correspond to the three scaled value dimensions discussed. The verbs are scored by means of an interval ranging from 1(1) to 7. Obviously, the values given need not be commented on, because they seem to covariate with common opinion. Thus in assigning values to the constructed concepts (cf. example 2) a covariation schedule can be set up which makes possible the study of the discriminating power in the valuation.

3.3 Representation of Discriminant Functions

Symbolic information encoded in verbal form is usually analyzed by means of a set of topics (or categories). The assumption underlying this approach is that the information structure is reflected in the structure of the category set. However, this kind of analysis very often proves to be insensitive to the concepts and conceptualizations that characterize a certain text. Furthermore, the manifest structure, especially of an interview text, does not allow for immediate inference. Based on this discussion the following hypothesis can be formulated:

Hypothesis 2:

In a multivariate study, by means of a discriminant analysis, the interaction between concepts and their scaled values can be used as a

basis for the identification of the structure characterizing symbolic information as defined by discriminant functions.

The extent to which the derived concepts can be used to study symbolic information multivariately has been investigated by means of the discriminant analysis procedure illustrated in Figure 2. The resulting discriminant functions are graphically presented in Figure 3, together with their standardized coefficients.

The process in Figure 2 starts with the coded text according to the AaO model from which A-O pairs are extracted on the basis of their coincidences as defined by the verb connecting them. Thereafter the process proceeds with the described cluster analysis.

The derived clusters define the score matrix ($p \times N$), that contains the values that the verbs have assumed through the scaling procedure. The

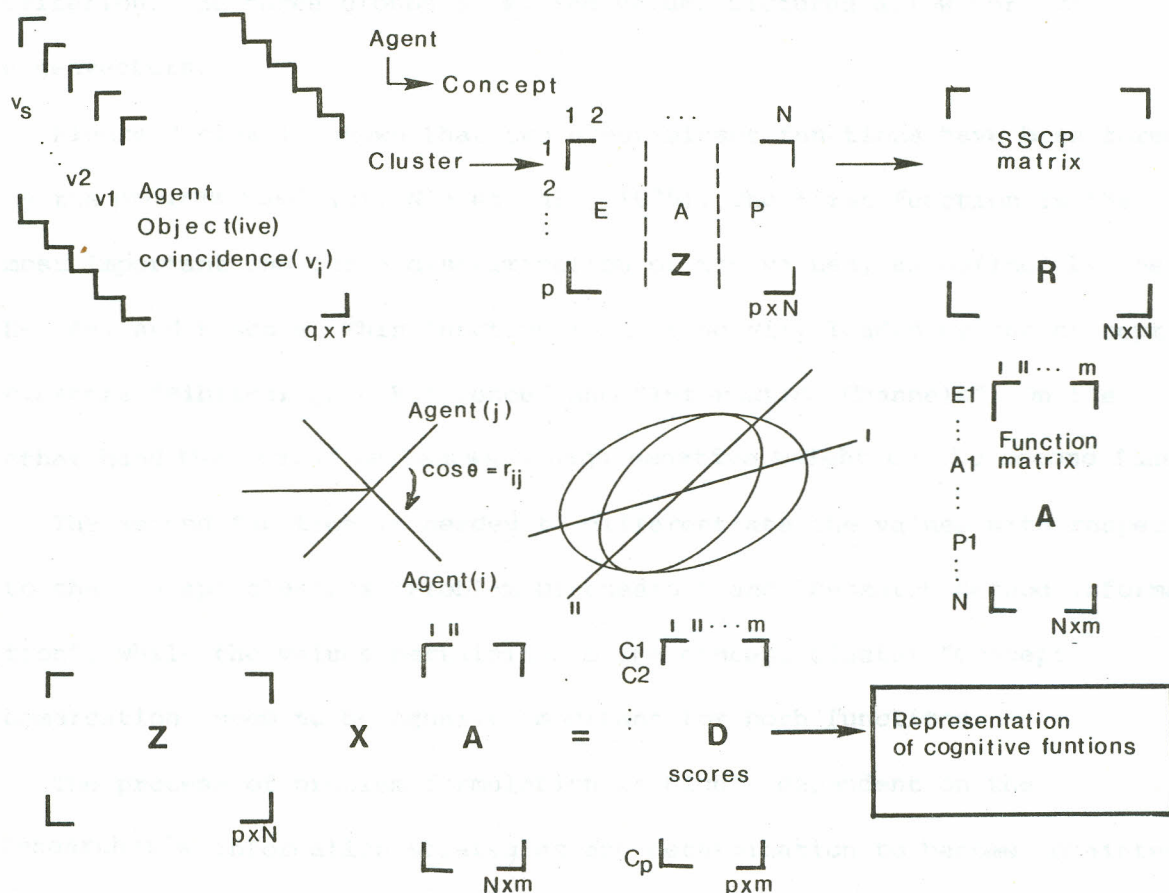


Figure 2. Multivariate Analysis of Running Text

concepts represent the predictor variables, i.e. 6 concept clusters ($p = 6$), while N stands for the total sample size ($N = 42$) comprising the three groups formed on the basis of the scaled value dimensions (E, A, P). From this a correlation matrix ($p \times p$) is then computed.

From the computed correlation matrix two $p \times p$ matrices can be constructed, where the first one expresses the within-group variance (W), while the other gives the between-group variance (B). On the basis of the within- and the between-group variances or correlations, eigenvalues are computed ($\lambda = W^{-1}B$). This expression indicates the discriminating power in the values as expressed by the corresponding linear combination of the values (see Tatsuoka, 1971).

The process of arriving at the first eigenvector value is pictured by the matrix equation ($D=ZA$), which provides a set of weights such that the resulting linear combination has the largest possible discriminant criterion. The three groups of scaled values pictured allow for two eigenvectors.

Figure 3 clearly shows that two discriminant functions have been formed by the program used (cf. Nie et al., 1975). The first function is the most important one for a discrimination of the values, as defined by the E-, A-, and P-scale. This function is most heavily loaded by the concept clusters "Bibliographic Reference" and "Information Channels". On the other hand the second one shows a high negative weight on the second function.

The second function is needed to differentiate the values with respect to the concept clusters "Problem Discussion" and "Research Method Information", while the values pertaining to the concept cluster "Concept Demarcation" seem to be equally important for both functions.

The process of problem formulation is highly dependent on the researcher's information strategies and determination to become acquainted

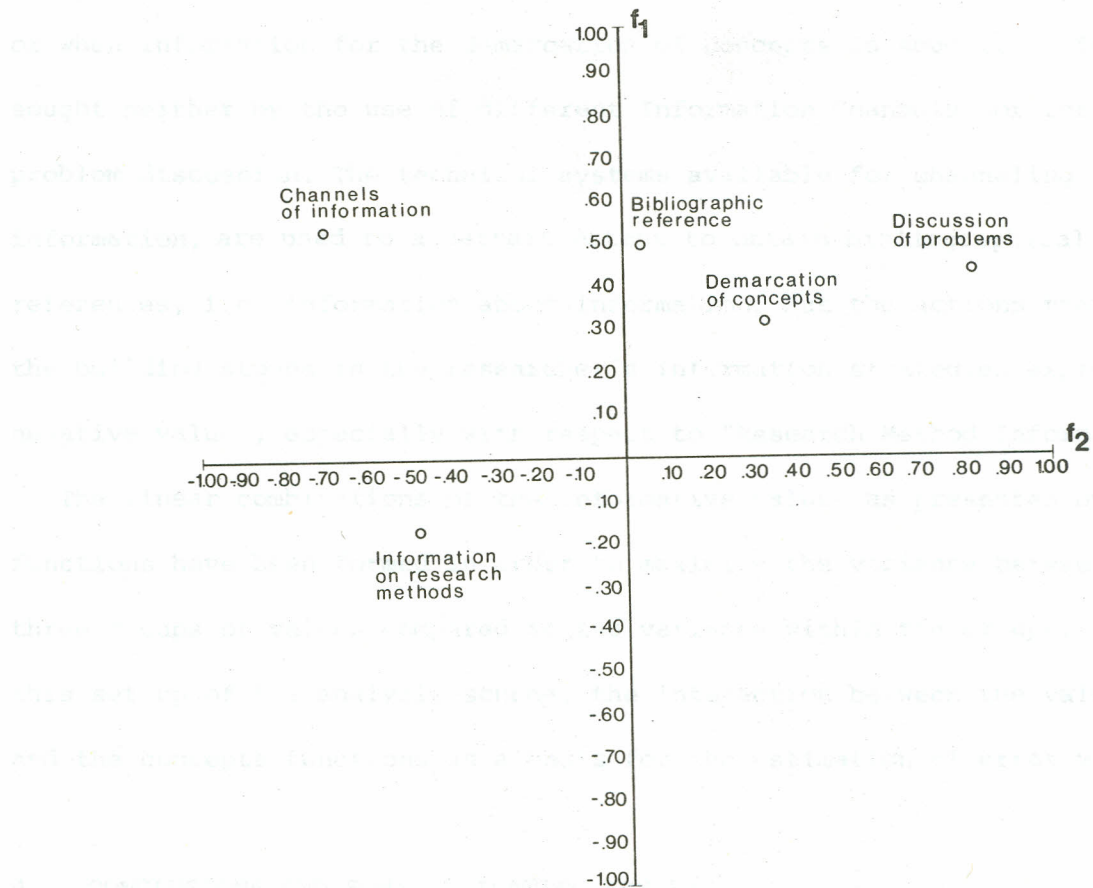


Figure 3. Discriminant Functions (I, II) and their Standardized Coefficients

with existing research information. His plan for solving his information problem contains intentions and goal notions, plus an idea of which means can be used to achieve the goals, i.e. means-goal hierarchies. As shown in Figure 3, his intention is to get in principle two types of information, (1) for demarcation of concepts, and (2) about research methods.

The strategy (means) that has been designed for the first type is problem discussion (discussion seminars, project meetings, informal literature seminars). The strategy designed for the second type is to a certain extent bibliographical information-seeking and visits to international symposiums, i.e. "Information Channels" are used.

Information strategies with respect to "Research Method Information" do not seem to be of any import while problem discussion is underway or when information for the demarcation of concepts is sought. It is sought neither by the use of different Information Channels nor through problem discussion. The technical systems available for channeling information, are used to a certain extent to obtain bibliographical references, i.e. information about information. But the actions that form the building stones in the researcher's information strategies express negative values, especially with respect to "Research Method Information".

The linear combinations of the informative values as presented by both functions have been formed in order to maximize the variance between the three groups of values compared to the variance within the groups. In this set up of the analysis scheme, the interaction between the values and the concepts functions as a basis for the estimation of error variance.

4. CONCLUSIONS AND SOME VALIDATING REMARKS

The ecological approach to the processing of symbolic information presented in this article constitutes an attempt to objectify originally subjective functions, namely the abstraction of information from anamorphic text.

The attempt demands that careful attention be directed to the circumstance that any presentation of something meaningful implies intentionality and the imposing of a perspective. An analysis that firmly incorporates them in its procedures seems to be able to extract the existing key relations.

Attempts have been made to assess information structures existing in verbal expressions by means of word associations. This approach is based on the assumption that the "similarity" between two linguistic elements

can be expressed as a relation between intersection and union of the distribution of these two elements.

An application of the word association model for the purpose of abstracting symbolic information from verbal expressions has been attempted and led to the following conclusions:

Conclusion 1:

Word association models lead to fairly rudimentary and therefore non-representative information structures.

The conclusion is based on the fact that words taken out of a running text do not reflect any valid information structure of that text. It was attempted to sort the agents into groups so that the degree of natural association is high between the agents that have been placed in the same group and low between the members of different groups. The same procedure was tried out on the object(ives). The analysis results suggest that this is not the right procedure. If the structural relation between an agent-object(ive) pair is used, it becomes easier for natural clusters appearing in the text to be formed.

Conclusion 2:

An ecological information processing model founded on the schema hypothesis leads to the abstraction of functions which seem to reflect the key relation underlying a running text.

The conclusion is based on the fact that the results reported from the discriminant analysis have been validated through a thorough and systematic check of the interview material. The validity has been demonstrated partly on the basis of an impressionistic analysis as reported in Annerblom (1974) and the questionnaire responses to seven-point bipolar scales presented in B. Bierschenk (1974).

The impressionistic analysis shows that the library is in many cases used because of good personal contact with the library staff. (N.B. the

cluster demarcating the concept "Information Channels" contains "person" as an element.) Reviewing organs such as ERIC and PA form one group that is used, but the researchers "feel dissatisfaction". Symposiums are attended roughly once a year, but only by senior staff members. Their evaluation shows that the information obtained is not valued very high.

Problem discussions seem to be the main source of ideas. Ideas and suggestions from single persons, obtained through face-to-face contact, telephone or letter are sought roughly once a month, which in this context refers to quite active behavior. This type of personal contacts is felt to be the best source of meaningful information on e.g. research methods.

Contrary to personal contacts, formal meetings and research seminars do not appear to provide valuable information. The concept cluster "Research Organization" covering these has had no discriminating power either, but rather a reducing effect when it is combined with the others. This further underlines the fact that their influence is regarded to be rather low.

These results support what has been abstracted from the symbolic information contained in the interview text. All the critical opinions on seminar forms and printed material indicate that "Information Dissemination" by the defined "Information Channels" does not well distribute information that has a strong influence on the formulation and demarcation of ill defined problems.

In the suggestions for improvement a desire is expressed for a better overall view and help in structuring the enormous flow of information. But it is also said that the researchers need not struggle through masses of literature for fear of missing something. Therefore, as is shown in Figure 3, the expected impact of information from these channels is considered to be rather low. These validating remarks complete the proof of the validity of the method presented.

5. REFERENCES

- Annerblom, M.-L. En impressionistisk innehållsanalys av intervjuer med forskare på pedagogiska institutioner i Sverige. /Interviews with researchers in departments of education in Sweden: An impressionistic analysis./ Pedagogisk-psykologiska problem, No. 255, 1974. /In Swedish/
- Bierschenk, B. Perception, strukturering och precisering av pedagogiska och psykologiska forskningsproblem på pedagogiska institutioner i Sverige. /The perception, structuring and definition of educational and psychological research problems at the departments of education in Sweden./ Pedagogisk-psykologiska problem, No. 254, 1974. /In Swedish/
- Bierschenk, B. A computer-based content analysis of interview texts. Numeric description and multivariate analysis. Didakometry, No. 53, 1977.
- Bierschenk, B. Conceptions of cognitive functions in a science of knowing. Mimeographed. Malmö: University of Lund, Department of Educational and Psychological Research, 1980.
- Bierschenk, B. & Bierschenk, I. A computer-based content analysis of interview data. (Studia Psychologica et Paedagogica, No. 32.) Lund: Gleerup, 1976.
- Bierschenk, I. Datorbaserad innehållsanalys: Kodningsmanual. /Computer-based content analysis: Coding manual./ Pedagogisk dokumentation, No. 52, 1977. /In Swedish/
- Broadbent, D.E. & Broadbent, M.H.P. The allocation of descriptor terms by individuals in a simulated retrieval system. Ergonomics, 1978, 21, 343-354.
- Cerccone, N. The representation and use of knowledge in an associative network for the automatic comprehension of natural language. Mimeographed. Burnaby, B.C.: Simon Fraser University, Computing Science Department, 1978.

- Di Lollo, V. Temporal integration in visual memory. Journal of Experimental Psychology: General, 1980, 109, 75-97.
- Fillmore, C.J. The case for case. In: E. Bach & R.T. Harms (Eds.). Universals in linguistic theory. New York: Holt, 1968.
- Gibson, J.J. The senses considered as perceptual systems. Boston: Houghton Mifflin, 1966.
- Gibson, J.J. The ecological approach to visual perception. Boston: Houghton Mifflin, 1979.
- Hubel, D.H. & Wiesel, T.N. Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. Journal of Neurophysiology, 1963, 26, 994-1002.
- Jenkins, J.J. Remember that old theory of memory? Well, forget it! In: R. Shaw & J. Bransford (Eds.). Perceiving, acting, and knowing: Toward an ecological psychology. Hillsdale, N.J.: Lawrence Erlbaum, 1977. Pp. 413-429.
- Kennedy, J.M. Pictures and the blind. Journal of the University Film Association, 1980, 32, 11-21.
- Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K. & Bent, D.H. Statistical package for the social sciences, (2nd ed.) New York: McGraw-Hill, 1975.
- Osgood, C.E. On the whys and wherefores of E, P, and A. Journal of Personality and Social Psychology, 1969, 12, 194-199.
- Paivio, A.U. Imagery and verbal processes. New York: Holt, 1971.
- Paivio, A., Philipchalk, R. & Rowe, E. Free and serial recall of pictures, sounds and words. Memory & Cognition, 1975, 3, 586-590.
- Piaget, J. Behavior and evolution. New York: Pantheon Books, 1978.

- Pylyshyn, Z.W. What the mind's eye tells the mind's brain: A critique of mental imagery. Psychological Bulletin, 1973, 80, 1-23.
- Quillian, M.R. Semantic memory. In: M.L. Minsky (Ed.). Semantic information processing. Cambridge, Mass.: M.I.T. Press, 1968. Pp. 216-270.
- Regan, D. & Beverley, K.I. Looming detectors in the human visual pathway. Vision Research, 1977, 18, 415-421.
- Regan, D., Beverley, K. & Cynader, M. The visual perception of motion in depth. Scientific American, 1979, 241, 136-151.
- Schank, R.C. Conceptual dependency: A theory of natural language understanding. Cognitive Psychology, 1972, 3, 552-631.
- Schank, R.C. Identification of conceptualizations underlying natural language. In: Schank, R.C. & Colby, K.M. (Eds.). Computer models of thought and language. San Francisco: Freeman, 1973. Pp. 187-247.
- Schank, R.C. The role of memory in language processing. In: C. Cofer (Ed.). The structure of human memory. San Francisco: Freeman, 1975. Pp. 162-189.
- Schank, R.C. & Abelson, R.P. Scripts, plans, goals and understanding. Hillsdale, N.J.: Erlbaum, 1977.
- Sperling, G. A model for visual memory tasks. Human Factors, 1963, 5, 19-31.
- Tatsuoka, M.M. Multivariate analysis: Techniques for educational and psychological research. New York: Wiley, 1971.
- Tulving, E. Episodic and semantic memory. In: E. Tulving & W. Donaldson (Eds.). Organization of memory. New York: Academic Press, 1972. Pp. 382-403.
- Uttal, W.R. The psychobiology of sensory coding. New York: Harper & Row, 1973.
- Werner, H. & Kaplan, B. Symbol formation. An organismic-developmental approach to language and the expression of thought. New York: Wiley, 1963.

Wickelgren, W.A. Age and storage dynamics in continuous recognition memory. Developmental Psychology, 1975, 11, 165-169.

Winston, P.H. (Ed.). The psychology of computer vision. New York: McGraw-Hill, 1975.

Wolsky, M. & Wolsky, A. The mechanism of evolution. A new look at old ideas. Basel: Kruger, 1976.

Wundt, W. Grundriss der Psychologie. Leipzig: Alfred Kröner Verlag, 1918.

Abstract card

Refecence card

Bierschenk, B. An ecological approach to the processing of symbolic information. Didakometry (Malmö, Sweden: Department of Educational and Psychological Research), No. 64, 1981.

This article attempts to prove that an adequate cognition oriented approach to the processing of symbolic information abstracted from verbal expressions must consider a running text instead of scattered sentences. It is argued that the dominant cognition models do not give due consideration to the fact that a valid abstraction of information structures has to be based on an explicit encoding of intentionality and valuation. Following a critical discussion of the foundations of the existing models, it is stated that a model is required that can cope with empirical context and novelty instead of truth-values in semantic-logical contexts. The proposed theory has been validated on empirical text material considered to be anamorphic. It is shown that its underlying key relations can be extracted by means of discriminant functions and given a meaningful interpretation.

Keywords: Abstraction, anamorphic text, conceptual functions, contextual information, perspective, symbolic information processing, specification, structural analysis.

Bierschenk, B. An ecological approach to the processing of symbolic information. Didakometry (Malmö, Sweden: Department of Educational and Psychological Research), No. 64, 1981.

ISSN 0046-0230